JPMTR-2219
DOI 10.14622/JPMTR-2219
UDC 394:004.89:159.937-021.415

Research paper | 174
Received: 2022-09-08
Accepted: 2023-01-11

Real-world evaluation of artificial intelligence-based color corrections for social media content creators

D. Hudson Smith and Erica B. Walker

Department of Graphic Communication, Clemson University,
207 Godfrey Hall, Clemson, SC, 29634,
United States of America

dane2@clemson.edu
eblack4@clemson.edu

Abstract

Brands strive to maintain consistent brand color representation across many digital channels, including social media platforms. This is a challenging goal given the varying real-world circumstances in which brand imagery is produced and shared. The ColorNet neural network tool was developed to automatically target and correct brand colors in imagery without altering non-brand colors. Previously, it was successfully applied to live sports broadcast footage. An open question is whether ColorNet can improve the accuracy, as measured by ΔE_{00}, of brand color representations in still photographs and photographs taken from videos and gifs posted to social media platforms such as Instagram and Twitter. To test this question, we collected a set of posts containing imagery from social media created by Clemson Athletics' social media accounts. We corrected the representation of Clemson orange in these images using ColorNet. After selecting pixel values corresponding to brand colors in each media piece, we demonstrate that ColorNet improves Clemson brand color accuracy across social media channels and for various media characteristics. Despite our observation that brand color representation varies significantly across media types and lighting conditions, the improvement in color representation from ColorNet was relatively consistent. We also showed that ColorNet has a comparatively minor impact on the color representation of skin tones.

Keywords: Twitter, Instagram, ColorNet, artificial neural networks

1. Introduction and background

Achieving consistent brand color is important to brands but challenging in real-world scenarios. Brands spend a significant amount of time determining and protecting brand specifications, including recognizable elements such as fonts, logos, and colors (Moser, 2003; Chang and Lin, 2010). However, consumers look at branded content across many different devices and channels and expect to see similar brand representation (Budelmann, Kim and Wozniak, 2010; Chang and Lin, 2010). For example, a branded product such as Coca-Cola is represented through physical packaging on the grocery shelf, on a jumbotron displaying a commercial during a live sporting event, and on the user's phone in an Instagram advertisement (Mayes, et al., 2021). Even if the actual product that appears on the television and phone screen is printed the correct color of red, the screens may not be color calibrated to display that same branded experience with color-accurate results (Conti and Walker, 2019).

ColorNet is a patented artificial intelligence algorithm that successfully detects and corrects a specified brand color in a live video feed (Walker, et al., 2020a; 2020b). ColorNet was initially developed to address brand color consistency on live video at sporting events. This study looks at the impact that ColorNet can have on content appearing across multiple branded social media platforms and asks the question:

Can ColorNet also improve color consistency on real-world social media content across platforms and content types?

During this study, another question arose when we noticed that the ColorNet algorithm impacted both the branded color (Clemson orange) and some skin tones. This discovery resulted in an additional set of data and analysis to better understand this undesired outcome. By determining situations where this happens, future development and training data will focus on minimizing and eliminating this unintended adjustment.

2. Methods

The Clemson Athletics Content Creation Team creates media for nineteen men's and women's National Collegiate Athletic Association (NCAA) sports, posting content for each across several social media platforms. Their main website highlights every team and includes direct links to an Instagram, Facebook, and Twitter feed with targeted content developed for each channel and sport.

For this study, we focused on content from Twitter and Instagram because both of those channels are used for different types of sports content, which would provide a wide variety of data types. Twitter is often used for play-by-play updates during the season, and Instagram focuses on more visual, summary-style content, including videos and photographs that have been post-processed after the event ends. We chose the following sports to analyze content from different lighting situations (indoor/outdoor, daytime/evening/nighttime): ClemsonTigers, ClemsonBaseball, ClemsonMBB (Men's Basketball), ClemsonWBB (Women's Basketball), ClemsonSoftball, ClemsonWSoccer (Women's Soccer), ClemsonWTennis (Women's Tennis), and ClemsonRowing.

The selection of media was randomized from an initial pool of 46 000 posts. Images that did not include any brand specification orange (for example, a graduate in the uniform of their new pro team or a student-athlete not wearing any branded clothing at a formal event) were removed from the analysis and replaced by the next randomized piece of content from the same channel and platform.

ColorNet 1.5 is a neural network-based model developed for automatic color correction on live video feeds (Mayer, Walker and Smith, 2021). We applied a version of this model pre-trained to correct Clemson orange to the full dataset of images collected from Instagram and Twitter. Original source content included photographs, vector-based graphics, animated gifs, and videos. A frame was automatically pulled from the moving content such as videos and gifs so that each source was represented by an individual jpg for this study. This resulted in a set of paired images where each pair contains the original jpg taken from social media and a version that was processed through ColorNet. Each pair of images was spatially aligned, allowing pixel-level comparisons before and after color correction.

We then labeled the original images to identify pixels that were either Clemson orange or skin tones using an online program called LabelBox, Figure 1. There was no specified limit or target number of pixels selected in each image. Instead, focus was placed on labeling a wide range of items in the image that were supposed to be branded Clemson orange. For example, brand logos found at the facility or on the scoreboard, uniform elements such as jerseys, helmets, shoes, and gloves, and graphic elements added during post-production. In addition, skin tone pixels were labeled with the intent of representing the widest possible range of skin tones across highlights, midtones, and shadows.

When labeling was complete for the full dataset, a JavaScript Object Notation (JSON) file was exported from LabelBox that was used to automate the measurements of each labeled pixel in red, green, and blue (RGB) color values before and after processing through ColorNet. We extracted the color values at each annotated pixel location for the image pairs, forming a set of corresponding color values. This set of paired color readings formed the basis for our analysis of the impact of ColorNet on color representation and consistency. For further analysis, we converted all RGB colors to CIELAB color space using the colormath Python library (Taylor, 2018).

For quantitative analysis, we used the ΔE_{00} (CIEDE2000, notated ΔE below for brevity) to characterize color representation for the points resulting from the above process. The ΔE measures a perceived visual difference between colors, and we were interested in measuring

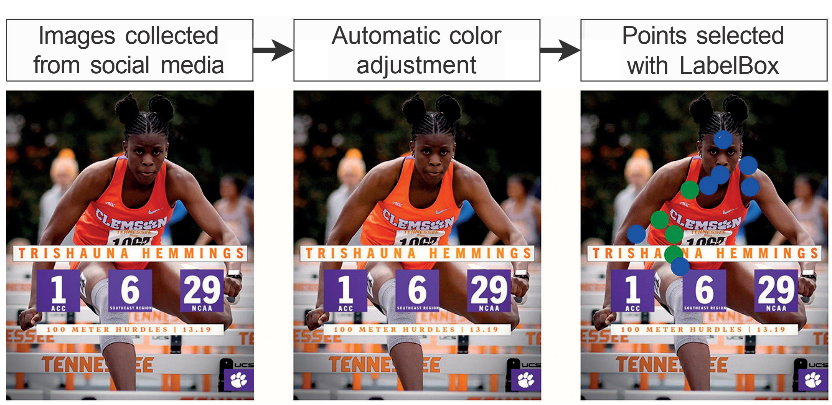
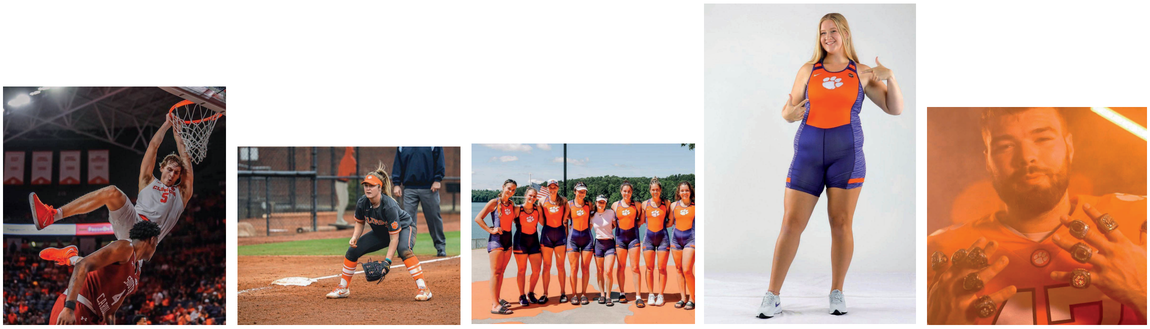


Figure 1: Data processing and annotation process



Bulb *Cloudy* *Sunny* *Studio* *Funky*

Figure 2: Examples of the five types of lighting scenarios

the difference in the readings, including the impact of luminance and saturation (Hunt and Pointer, 2011, pp. 61–68).

For brand color pixels, we measured the difference between the observed color and the official brand color specification for both the original and color-adjusted samples (Clemson orange brand, n.d.). To improve color representation, ColorNet should shift the distribution of ΔE values toward zero. In addition, for both brand color and skin tone points, we measured the ΔE between the original and adjusted imagery. We expected ColorNet to produce large ΔE values for the brand color points and small values for skin tones. After computing these ΔE values for each selected point, we examined the performance characteristics in aggregate, comparing across platforms and channels, lighting conditions, and type of source content (still or video).

Lighting was determined to belong to one of five categories that describe the type of lighting present during the initial capture through visual analysis: bulb, cloudy, sunny, studio, or funky (Figure 2). Bulb represented indoor or outdoor situations that were artificially lit. For example, night time on the soccer field with the stadium lights powered on or indoors in the basketball arena. Cloudy or sunny lighting both came from outdoor content. Sunny lighting produced harsh highlights and shadows, whereas cloudy lighting produced more even tones across the image. Content marked as studio lighting was taken, with additional professional lights, in a controlled indoor location. Funky lighting indicates content that was created with lights that have orange or purple gels that skew the actual color of the content in the images.

Qualitative analysis provided a way to more deeply understand which tones ColorNet adjusted more or less accurately. For this process, we looked at the thirty best and worst adjustments to the brand specification orange pixels and the skin tone pixels across both platforms. For brand orange, this was defined as the highest and lowest ΔE values between the corrected pixel and the target brand color. The skin tone pixels that were adjusted the least (closest to a zero ΔE) and those that were adjusted most significantly (largest ΔE value) from the original skin tone measurement were also examined. Analysis included looking for trends across accounts, platforms, content types (photo, video, or gif), and lighting types (bulb, cloudy, sunny, studio, or funky).

We collected a total of 46 000 messages from nine Twitter and nine Instagram accounts. We then randomly sampled 90 messages containing photographs, graphics, videos, or gifs, 10 from each account, and collected still-images for each post. Posts that included moving images (gif or videos) grabbed the platform-provided thumbnail image from the content. For brevity we refer to such thumbnails as “gif” or “video.” Across these 180 images (90 from Twitter, 90 from Instagram), we annotated a total 2 457 points as examples of brand color representation (1129 points) or skin tone (1328). Each image was also labeled according to the lighting condition and content type. The number of images representing each category and other details about the collected data set are summarized in Table 1.

Table 1: Summary of the collected data

		Twitter	**Instagram**	**Total**
	# accounts	9	9	18
# images	photo	77	51	128
	video	11	39	50
	gif	2	0	2
	total	**90**	**90**	**180**
# lighting	bulb	27	38	65
	sunny	23	22	45
	studio	17	14	31
	cloudy	11	12	23
	funky	8	3	11
	N/A	4	1	5
	total	**90**	**90**	**180**
# points	brand color	537	592	1 129
	skin tone	700	628	1 328
	total	**1 237**	**1 220**	**2 457**

3. Results

3.1 Measurement of color fidelity

We measured the difference between the 1129 selected brand color points and Clemson orange brand specification (denoted by the brand as RGB 245, 102, 0 and Pantone 165, Figure 3). The CIELAB values are not provided by the brand so RGB color readings were converted to CIELAB and back to RGB using the colormath Python library for processing and analysis. Before color correction, the average ΔE value was 14.9 ± 0.6. After color correction, the average ΔE improved to 11.5 ± 0.5. While these values are still large, measurement against brand specification is too strict as it neglects appropriate variation in color presentation due to lighting conditions such as deep shadows or blown out highlights. The average ΔE in the 1.3k selected skin tone points between the original and corrected images was 2.6 ± 0.1. Though this is larger than ideal, it is encouraging that the shift is much smaller than the corresponding shift for the brand color points (7.1 ± 0.2). This suggests that ColorNet can differentiate between brand color regions and skin tone regions – a challenging task since the tonal range of skin falls close to Clemson brand orange in certain situations.

We next examined how brand color representation and ColorNet performance varies by platform, media type, and lighting condition. Figure 4a shows the mean performance and 95 % confidence intervals for both platforms for the original and corrected imagery. On average, media on the two platforms have similar brand-color fidelity; the improvements from ColorNet are consistent. When we consider lighting conditions (Figure 4b), we see that studio lighting conditions produce significantly better color fidelity in the original and corrected images. Looking at the ΔE values across media types (Figure 4c), we find the surprising result that still photographs have much better color fidelity than videos both before and after correction. We speculate that this may result from the different cameras and camera settings used by the media teams and post-processing techniques applied differently in the two cases.

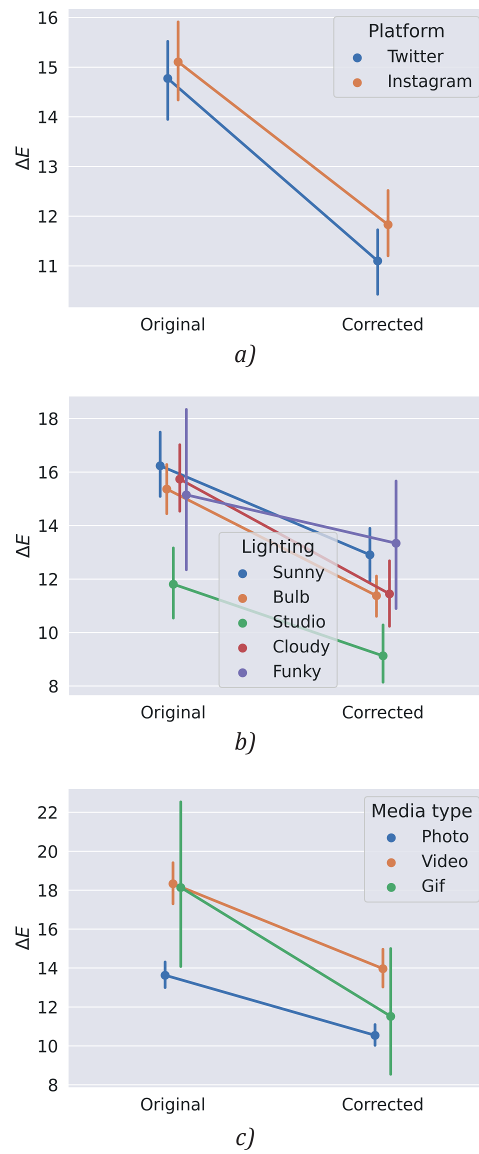


Figure 4: Performance by platform (a), lighting condition ('N/A' samples are omitted) (b), and media type (c); the Y-axes measure ΔE between the collected color points and Clemson brand orange, the "video" and "gif" categories are based on still images collected from videos or gifs

Figure 3: Brand designations across different media (Clemson orange brand, n.d.)

Finally, to better understand the average ΔE of 2.6 ± 0.1 for the 1.3k selected skin tone points, we evaluated the average shift by platform, lighting condition, and media type. We found little variation in the skin tone shift by platform or lighting condition. However, when considering media type, we observed an average shift of 2.8 ± 0.1 for photos and an average shift of 2.0 ± 0.2 for videos. In other words, ColorNet led to a greater shift in skin tones for photos than for video. This may be a consequence of the poorer color fidelity in the videos causing skin tones to shift further away from Clemson brand orange, as indicated in the previous paragraph, or it may indicate a deficiency in ColorNet itself.

3.2 Analysis of best and worst-performing pixels

For brand orange, the best-performing pixels were closest to brand specification after being corrected by the ColorNet algorithm. For the top 30 pixels, the ΔE values ranged from 0.49 to 1.69. The distribution included ten accounts with the top four represented by: ClemsonSoftball, ClemsonBaseball, ClemsonWBB, and ClemsonWSoccer. Twice as many came from Instagram as from Twitter (20, 10) despite a nearly even number of source points in the original data set (1237, 1220). Nearly twice as many were from photographic rather than video sources (19, 11), but it should be noted that the full data set includes more than three times the number of photograph content versus video so there may be greater diversity of photographic situations in the original data (1793, 637). Lighting was distributed across all five listed types with bulb and sunny having the most pixels in the top thirty best performers. Further qualitative analysis showed that ColorNet corrects Clemson orange best on uniforms in neutral lighting or highlights and in several cases did not manipulate parts of the stadium that should be brand orange but do not look to be close to specification due to color fade and sun exposure.

For brand orange, the worst-performing pixels were considered pixels that should appear as brand color but were furthest from brand specification after being corrected by the ColorNet algorithm. For the bottom 30 pixels, the ΔE values ranged from 34.3 to 44.4. It should be noted that the original values before processing with ColorNet ranged from 34.6 to 48.0, so these pixels were initially displaying far from the desired color specification even before being processed. ColorNet adjusted the pixels by an average of 5.26 ΔE. The distribution included eleven accounts with ClemsonSoftball and ClemsonBaseball again appearing as the most frequent accounts. Slightly fewer than twice as many came from Instagram as from Twitter (17, 13) and were from photographic rather than video sources (17, 13). Lighting was distributed across all five listed types, with sunny and bulb having the most pixels in the worst thirty performers. Further qualitative analysis showed that Clemson orange displays furthest from specification in shadowed or very shadowed areas of the frame or occasionally in very low-quality images such as a still image reproduced from a television broadcast.

For skin tones, we ideally want ColorNet to make no adjustments, but, since that was not what the data showed, qualitative analysis helped the researchers identify parts of the algorithm and training data that need further adjustment. The best-performing skin tone pixels were very minimally adjusted or not adjusted at all by the ColorNet algorithm. For the top 30 pixels, the ΔE values ranged from 0.0 to 0.31. The distribution included thirteen accounts with the top four: ClemsonBaseball, ClemsonWTennis, ClemsonWBB, and ClemsonRowing. The distribution between platforms was nearly even, with 16 from Twitter and 14 from Instagram. More were from photographic rather than video sources (18, 12). Lighting was distributed across all five listed types, with bulb lighting taking a heavy lead (17) and sunny in second (6). Further qualitative analysis showed that ColorNet does not tend to correct lighter skin tones in neutral or highlight areas of the frame, with only three out of the least affected skin tone samples including dark skin in neutral lighting.

The worst-performing skin tone pixels were those that were adjusted the most by the ColorNet algorithm. For these 30 pixels, the ΔE values ranged from 7.9 to 15.3, all adjustments that we would consider unacceptable for skin tone representation. The distribution included twelve accounts, with the top four most frequently appearing accounts as follows: ClemsonRowing, ClemsonBaseball, ClemsonTigers, and ClemsonFB. Pixels from Twitter content appeared twice as often as from Instagram (20, 10). Most of the content was from photographs rather than videos (27, 3). Lighting was distributed across all five listed types, with sunny and cloudy lighting conditions appearing most frequently (14, 8). Further qualitative analysis showed that ColorNet can incorrectly identify dark skin tones in neutral lighting and suntanned lighter skin tones in shadow as intended to be brand color. It was also noticed that skin on ears and around the lips are more often confused and therefore more frequently adjusted by ColorNet than other areas of skin.

4. Discussion

Initially, a set of images from social media posts by Clemson Athletics marketing accounts was pulled, paired, and corrected using the ColorNet neural network to address brand color consistency. Relative to the imagery taken directly from social media, ColorNet produces brand color representations with smaller deviations, on average, from true Clemson orange resulting in a positive improvement in brand color representation. This pattern holds when considering multiple channels and media characteristics in aggregate and for each stratum individually. Taken at face value, this suggests that the ColorNet neural network, originally developed for broadcast footage, can also improve brand color fidelity for use with social media content when applied to Clemson brand colors. This quantitative approach assumes that low deviations from brand color specification correspond to a positive value in the eyes of the brand owners and fans. Taken to an extreme, however, this is clearly incorrect, as a distribution with no deviation around brand specification would lose natural variations in the presentation of color caused by natural highlights and shadows present in the environment. Future user studies are needed to assess whether marketers and content creators would value the ability to produce more accurate brand representations. Should that be desired by the brand stakeholders, further development of an integrated, quick turnaround tool would be necessary for full integration and implementation at the point of content creation and distribution.

5. Conclusions

Through this study, we demonstrated that ColorNet improves brand color fidelity across social media channels and for various media characteristics for Clemson brand colors. Despite our observation that brand color representation varies significantly across media types and lighting conditions, the improvement of color representation after processing media through ColorNet was relatively consistent. This study also showed that ColorNet has a comparatively minor impact on the color representation of skin tones in imagery but further refinement of ColorNet to address this would be a valuable undertaking to ensure the lowest impact possible on skin tone areas present in the image.

Acknowledgments

The researchers would like to acknowledge the Powers College of Business summer research program, the Creative Inquiry (CI) program, and The Robert H. Brooks Sports Science Institute.

References

Clemson orange brand, n.d. *Bringing Clemson into hue – color palette.* [online] Available at: <https://www.clemson.edu/brand/guide/color.html> [Accessed 1 September 2022].

Budelmann, K., Kim, Y. and Wozniak, C., 2010. *Brand identity essentials: 100 principles for designing logos and building brands*. Beverly, MA, USA: Rockport Publishers.

Chang, W.-L. and Lin, H.-L., 2010. The impact of color traits on corporate branding. *African Journal of Business Management*, 4(15), pp. 3344–3345.

Conti, B. and Walker, E.B., 2019. Are you for the orange team or the red team? An exploratory study on displaying brand colors on large format LED screens at live sporting events. *Visual Communications Journal*, 55(1), pp. 3–8.

Hunt, R.W.G. and Pointer, M.R., 2011. *Measuring colour*. 4th ed. Chichester, UK: John Wiley & Sons.

Mayer, M., Walker, E.B., Smith, D.H., 2021. ColorNet 1.5: Improving machine learning color correction for video. In: *Proceedings of the 2021 NAB Broadcast Engineering and IT Conference.* Online, 9–13 October 2021. Washington, DC, USA: National Association of Broadcasters.

Mayes, E., Lineberger, J.P., Mayer, M., Sanborn, A., Smith, H. and Walker, E., 2021. Automated brand color accuracy for realtime video. *SMPTE Motion Imaging Journal*, 130(3), pp. 45–49. https://doi.org/10.5594/JMI.2021.3058397.

Moser, M., 2003. *United we brand: how to create a cohesive brand that's seen, heard, and remembered*. Boston, MA, USA: Harvard Business School Press.

Taylor, G., 2018. *Colormath 3.0.0*. [online] Available at: <https://pypi.org/project/colormath/3.0.0/> [Accessed August 2022].

Walker, E.B., Smith, D.H., Lineberger, J.P., Mayer, M.L., Mayes, E.E. and Sanborn, A.T., 2020a. ColorNet: a neural network-based system for consistent display of brand colors for video. In: *Proceedings of Society for Information Display 2020 International Symposium*, 51(1). Virtual, 3–7 August 2020. Campbell, CA, USA: Society for Information Display.

Walker, E.B., Smith, D.H., Mayes, E.E., Lineberger, J.P., Mayer, M.L. and Sanborn, A.T., 2020b. Consistent display of Clemson brand colors using artificial intelligence. In: *TAGA Proceedings: 72nd Annual Technical Conference*. Oklahoma City, OK, USA, 15–18 March 2020. TAGA.